一种基于 DTMF 信号的智能手机外部攻击方法 *

申兴发，杨　健†，冉德纲

(杭州电子科技大学 计算机学院，杭州 310018)

摘　要：传统的针对智能手机内部攻击的方式容易被用户察觉及预防。作为一种常见的音频信号，DTMF 信号在手机通信中具有非常重要的地位，但也面临严峻的安全风险。提出了一种基于 DTMF 信号的智能手机外部攻击方法，可以在用户不被察觉，且与用户手机无交互情况下进行有效攻击。首先，该方法对用户某些重要按键操作进行录音；然后对录音数据在时域上进行双阈值的端点检测，提取信号的有效区域；再将有效区域通过 Goertzel 算法转换到频域进行数字分类；最后，通过比照 DTMF 编码表得到用户所有按键数据。实验结果表明，该方法在 10 db 信噪比，且与用户手机无交互的条件下能破解 80%以上的按键数据。

关键词：智能手机；内部攻击；DTMF 信号；外部攻击；端点检测；Goertzel 算法

中图分类号：TN918.91　　doi: 10.19734/j.issn.1001-3695.2018.11.0893

External attack method for smartphone based on DTMF signal

Shen Xingfa, Yang Jian†, Ran Degang

(*School of Computer Science & Technology*, *Hangzhou Dianzi University*, *Hangzhou* 310018, *China*)

Abstract: The traditional way of internal attack for smartphone is easy to detected and prevented by the user. As a common audio signal, DTMF signal plays a very important role in mobile communication, but also faces severe security risk. This paper proposed an external attack method for smartphone based on DTMF signal, which could attack effectively without the user being aware and without interaction with the smartphone. Firstly, it recorded some important keystroke operation of user. Secondly, performed double-threshold endpoint detection in time domain to extract the effective area of the signal. Thirdly, converted the effective area to frequency domain by Goertzel algorithm for digital classification. Finally, all the keystroke data of the user were obtained by comparing the DTMF coding table. The experimental results show that the method can decipher more than 80% of the keystroke data under the condition of 10db signal-to-noise ratio and no interaction with the smartphone.

Key words: smartphone; internal attack; DTMF signal; external attack; endpoint detection; Goertzel algorithm

0　引言

随着计算机技术和通信业的发展，使用智能手机的人越来越多。根据互联网分析师玛丽•米克尔最新互联网趋势报告[1]显示，到 2018 年，全球智能手机的用户人数约为 15 亿，中国智能手机用户数量已达 3.54 亿，超越美国成为世界上智能手机用户量最多的国家。移动智能手机的广泛使用使它们成为进行隐私和安全攻击的目标，各种针对智能手机的攻击接踵而至，其中有三类攻击最为频繁，第一类是软件层面的攻击，如 Das 等人发现的一些第三方应用可以收集私自用户物理位置数据以及联系人信息，Schlegel 等人[3]发现的一些恶意软件可以捕捉语音通话和记录内置麦克风的任何对话；第二类是系统层面的攻击，如 Drake 等人[4]发现了 Android 系统上多种远程代码执行漏洞，并且不需要用户交互就可以进行攻击，Xing 等人[5]发现的针对 iOS 操作系统缺乏身份验证造成网络套接字滥用问题；第三类网络层面是的攻击，如 Manikandan 等人[6]提出的通过移动热点通信信道获取其他节点的数据和流量，Cheng 等人[7]提出的公共 WiFi 接入点的隐私泄露问题。这三类针对智能手机的攻击都属于内部攻击的方式，都容易被用户察觉并采取相应的解决方法，如 Akhuseyinoglu 等人[8]针对软件层面的攻击提出了一种基于机器学习的恶意软件自动检测工具。Enck 等人[9]针对系统层面的攻击提出的 TaintDroid 系统执行动态污染分析，可以显示智能手机系统中隐藏的和可能不需要的私有数据信息流。Khoula 等人[10]针对网络层面的攻击提出了一种针对移动热点入侵检测系统 SIDS。

综上所述，已有文献的研究成果主要关注的是对智能手机内部的攻击，且用户容易发现并采取相应措施进行解决，反而忽视了对智能手机外部进行攻击的方式。由于智能手机具有体积小，功能强大等特点，大多数的人都选择智能手机作为业务工具，像手机银行、手机邮件、人工服务和手机支付等等，然而这些业务操作中都会使用智能手机的拨号键盘，用户通过拨号键盘进行菜单选择以及密码输入，而这种按键操作会产生一种 DTMF 的音频信号，通过这种 DTMF 信号的录音进行检测可以识别用户的按键内容，且不容易让用户察觉，也不需要与用户手机进行交互。本文基于此考虑，提出了一种基于 DTMF 信号智能手机外部的攻击方法，攻击者通过对受害者某种操作(密码输入、电话号码输入等)按键音

收稿日期：2018-11-28；**修回日期：**2019-01-16　**基金项目：**国家自然科学基金资助项目（61672198）；杭州电子科技大学学科交叉创新团队建设项目

作者简介：申兴发（1979-），男，安徽宣城人，教授，硕导，主要研究方向为物联网、移动计算、大数据、信息安全等；杨健（1994-），男(通信作者)，湖北黄冈人，硕士研究生，主要研究方向为语音信号处理、信息安全等（jYang94@foxmail.com）；冉德纲（1992-），男，山东德州人，硕士研究生，主要研究方向为物联网、嵌入式等.

的录音，然后在时域上通过双阈值端点检测算法将按键产生的连续 DTMF 信号分割成每一个按键的 DTMF 信号，在频域上将分割的每一个按键的 DTMF 信号通过 Goertzel 算法上解析成相应的数字，最后获取用户的密码和手机号等隐私数据。经证明，该攻击方式在 10 db 的信噪比的条件下，可以破解 80%以上的 4 或 6 位数字密码和 11 位数字的手机号码。

1　DTMF 信号

电话通信具有两种拨号方式，一种是脉冲式拨号，一种是 DTMF(dual tone multi frequency)，也就是双音多频[11,12]式拨号。脉冲式拨号是早期常用的拨号方式，它通过彼此断开的脉冲序列发送呼叫和被叫号码。但随着使用者和使用频率的增加，脉冲式拨号暴露出人工量大，误码率高，通信效率低的问题。因此美国贝尔实验室于 1963 年推出了一种双音多频式拨号方式，以准确的、高效的完成电话通信网络中被叫号码的自动接收，随后双音多频电话迅速取代了脉冲式电话成为主流的拨号方式。

如今，DTMF 信号已经成为电话系统、智能监控系统以及远程控制系统的主要通信方式。如表 1 所示，每一个 DTMF 信号由两个频率的音频信号叠加构成，这两个音频信号的频率来自两个预分配的频率组：行频组和列频组。每一对这样的音频信号唯一产生表示一个数字或符号。DTMF 编码是将键盘上不同的按键用 8 个频率的不同组合表示出来，如当人们在智能手机的拨号键盘按下数字“1”时，将会由 697 Hz 和 1209 Hz 的音调进行叠加而来产生数字“1”的 DTMF 信号。

表 1　DTMF 信号编码表

Table 1　The coding table of DTMF signal

Hz	1209	1336	1477	1633
697	1	2	3	A
770	4	5	6	B
852	7	8	9	C
941	*	0	#	D

2　基于时域的端点检测

考虑到用户在智能手机上输入密码或电话号码的动作是连续的，大部分的密码是 4 位或者 6 位，手机号码是 11 位，所以攻击者录音时得到的是连续的 DTMF 信号，需要将连续的 DTMF 信号分割成每一个按键的 DTMF 信号进行识别。因此本文将语音信号识别中的端点检测[13,14]技术引入到 DTMF 信号检测，从连续的 DTMF 信号中分割出按键信号和非按键信号。同时用户的按键内容是随机的，可能会出现一些重复性的数字或字母。例如用户进行电话银行操作时，拨打“95588”这种类型的数字号码。通过对号码“95588”的 DTMF 信号进行时域和频域上的分析，如图 1 所示，可以看出相同的数字在频域上表现出相同的频率增幅，并没有像时域上呈现多个独立的信号，对于 DTMF 信号的分割会造成很大误差。基于此，本文采用基于时域的端点检测算法区分 DTMF 信号和非 DTMF 信号。

2.1　短时能量

由于语音信号能量随时间变化而变化，语音段和噪声段之间的能量差别相当显著。因此可以通过对语音信号的短时能量进行分析，可以描述语音的这种特征的变化。设语音时域信号为 $x(l)$，窗函数为 $w(m)$，加窗后的第 n 帧语音信号为 $x_n(m)$，则有

$$x_n(m) = w(m)x(n+m),\ 0 \le m \le N-1,\ \ n = 0,\ 1T,\ 2T,\ \cdots \quad (1)$$

其中：N 为帧长，即窗口的长度，T 为帧移。

加窗后的第 n 帧语音信号 $x_n(m)$ 的短时能量为

$$E_n = \sum_{m=0}^{N-1} x_n^2(m) \quad (2)$$

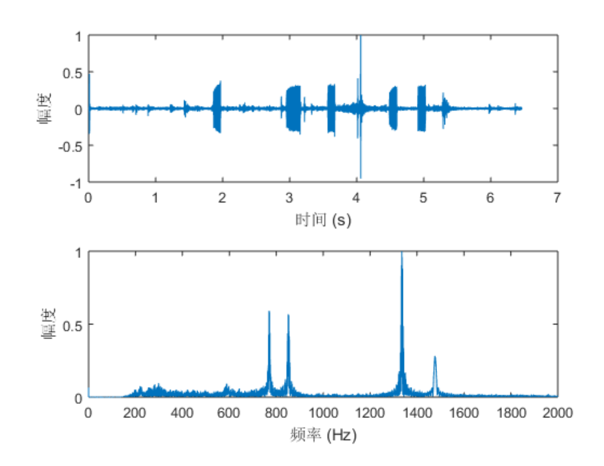


图 1　“95588”的 DTMF 信号在时域和频域上的分布

Fig. 1　Distribution of "95588" DTMF signals in time domain and frequency domain

2.2　短时过零率

时域分析中除了对语音信号的短时能量进行分析，还可以从语音信号波形的角度分析。语音信号输入会随着时间上下波动，形成其特有的波形，这一特征可以用一帧内语音信号通过零值得次数来描述，称为短时过零率。若是连续信号，则语音信号的时域波形通过横轴的次数即为过零；若是离散信号，则相邻取样值之间的符号变化即为过零。由于语音信号是一种短时平稳信号，采用短时过零率可以在一定程度上反映语音信号的频谱性质，由此可以获得频谱特性的一种估计。设语音信号为 $x(l)$，窗函数为 $w(m)$，加窗后第 n 帧语音信号 $x_n(m)$ 的短时过零率 Z_n 为

$$Z_n = \frac{1}{2}\sum_{m=1}^{N-1} \left| \mathrm{sgn}(x_n(m)) - \mathrm{sgn}(x_n(m-1)) \right| \quad (3)$$

其中 sgn() 是符号函数。

$$\mathrm{sgn}(x) = \begin{cases} 1,\ x \ge 0 \\ -1,\ x < 0 \end{cases} \quad (4)$$

2.3　双阈值端点检测

在语音信号信噪比高的条件下，仅靠计算信号的短时能量就基本能够把语音段和噪声段区别开来，即使在最低电平的语音的能量也比噪声能量要高。但是，在现实生活中很难保证信号的高信噪比，仅仅根据短时能量进行判断是不够的。因此，还需进一步利用信号的短时过零率进行判断，而语音段的短时平均过零率比噪声的短时平均过零率要高出好几倍，所以可以作为区分语音段的辅助判断。基于此，本文利用信号的短时能量和短时过零率双重阈值进行判别 DTMF 信号和非 DTMF 信号。

a)基于短时能量判别。如图 2 所示，首先根据语音信号的短时能量平均值选取一个较高的阈值 T_1，进行第一次粗判，语音的起止点位于该阈值与短时能量曲线交点所对应的时间间隔之外(即 A_1A_2 段之外)。再根据背景噪声的平均能量确定一个较低的阈值 T_2，并从 A_1 点往左、从 A_2 点往右进行搜索，分别找到短时能量曲线与 T_2 相交的两个点 B_1 和 B_2，于是 B_1B_2 就是基于短时能量所判断的语音段。

b)基于短时过零率判别。如图 2 所示，根据背景噪声的平均过零率确定一个阈值 T_3，以信号的短时平均过零率为标

准，从 B_1 点往左、从 B_2 点往右搜索，找到短时平均过零率低于阈值 T_3 的两个点 C_1 和 C_2，这就是语音段的起止点，即 DTMF 信号的起止点。

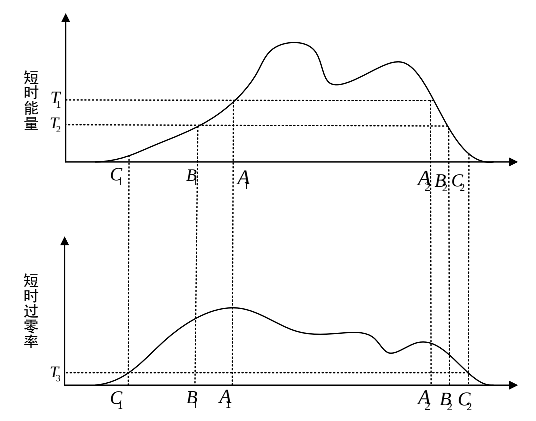


图 2 基于短时能量和短时过零率的双门限判别示意图

Fig. 2 Dual threshold discrimination diagram based on short-time energy and short-time zero-crossing rate

通过短时能量的粗判，在结合短时过零率的特性进行辅助判断，可以很好的分割 DTMF 信号的有效区域。如图 3 所示，数字“95588”的连续按键音信号经过双阈值端点分割后的情况，从上往下依次是数字“95588”录音的原始信号、在加入 10db 的噪声后的信号、信号的短时能量分布以及信号的短时过零率分布，从图中可以看出该双阈值端点检测算法能够很好的分辨出 DTMF 信号和非 DTMF 信号。

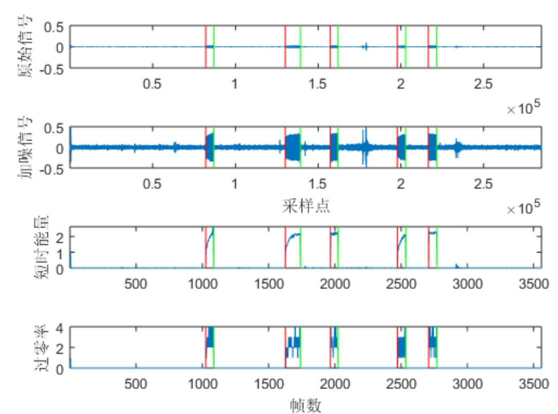


图 3 数字“95588”的端点分割情况

Fig. 3 Endpoint segmentation of the number "95588"

3 基于频域的分类检测

连续的 DTMF 信号经过时域上的端点检测后，被分割成单个的 DTMF 信号，需要对单个的 DTMF 信号进行分类解析成表 1 中的数字和字母。传统的 DTMF 信号检测方法有 DFT 算法、FFT 算法等，本文采用 Goertzel 算法对 DTMF 信号进行检测，相对于 FFT 算法效率更高。

3.1 Goertzel 算法原理

Goertzel 算法[15]是由美国人杰拉德 · 戈泽尔(Gerald Goertzel)在 1958 年提出，主要用于数字信号处理领域中，是一种计算信号离散傅里叶变换的方法。设语音信号为 $x(n)$，则信号第 k 个 DFT 分量表示为

$$X(k)=\sum_{n=0}^{N-1}x(n)e^{-j2\pi k\frac{n}{N}},\ k=0,\ \cdots,\ N-1 \tag{5}$$

又因为

$$e^{j2\pi k\frac{N}{N}}=1 \tag{6}$$

对等式(5)两边同时乘以 $e^{j2\pi k\frac{N}{N}}$ 可以得到

$$X(k)=e^{j2\pi k\frac{N}{N}}\sum_{n=0}^{N-1}x(n)e^{-j2\pi k\frac{n}{N}} \tag{7}$$

通过对等式(7)进行变形可以得到

$$X(k)=\sum_{n=0}^{N-1}x(n)e^{-j2\pi k\frac{n-N}{N}} \tag{8}$$

等式(8)的右边部分可以看成信号序列 $x(n)$ 与脉冲响应 $h_k(n)$ 的离散线性卷积。其中 $h_k(n)$ 可以表示为

$$h_k(n)=e^{j2\pi k\frac{n}{N}}u(n) \tag{9}$$

如果把卷积计算的结果表示为 $y_k(n)$，则有

$$y_k(n)=\sum_{m=0}^{N-1}x(m)e^{j2\pi k\frac{n-m}{N}}u(n-m) \tag{10}$$

由此，可以对输出响应进行 Z 变换得到变换方程：

$$H_k(z)=\sum_{n=-\infty}^{\infty}h_k(n)z^{-n}=(\sum_{n=0}^{\infty}e^{j2\pi\frac{k}{N}}z^{-1})^n \tag{11}$$

其中等式(11)可以看成以 $e^{j2\pi k\frac{0}{N}}z^{-0}=1$ 为首项，以 $e^{j2\pi\frac{k}{N}}z^{-1}$ 为公比的几何级数等比序列。对于 $|q|<1$，即 $|z|>1$ 的情况，此序列是趋于收敛的，且其求和结果为

$$H_k(z)=\frac{1}{1-e^{j2\pi\frac{k}{N}}z^{-1}} \tag{12}$$

其相应的差分方程为

$$y_k(n)=x(n)+e^{j2\pi\frac{k}{N}}y_k(n-1) \tag{13}$$

其中 $y_k(-1)=0$。

为了解决差分方程计算过程中带有复杂的倍增系数，可以在差分方程的分子和分母中同时引入共轭因子 $(1-e^{-j2\pi\frac{k}{N}}z^{-1})$，从而构造了一个含有二阶共轭极点的滤波器，进而可以得到

$$H_k(z)=\frac{1-e^{-j2\pi\frac{k}{N}}z^{-1}}{1-2\cos(\frac{2\pi k}{N})z^{-1}+z^{-2}} \tag{14}$$

其二阶差分方程为

$$y_k(n)=x(n)-x(n-1)e^{-j2\pi\frac{k}{N}}+2\cos(\frac{2\pi k}{N})y_k(n-1)-y_k(n-2) \tag{15}$$

其中 $x(-1)=y(-1)=y(-2)=0$。对上式引入中间变量 $v(n)$，可以得到最终的结果为

$$\begin{cases}v(n)=x(n)+2\cos(\frac{2\pi k}{N})v(n-1)-v(n-2)\\ y_k(n)=v(n)-e^{-j\frac{2\pi k}{N}}v(n-1)\end{cases} \tag{16}$$

该滤波器对应的网络结构如图 4 所示，其中 $x(n)$ 是语音信号采样值，$v(n)$ 是中间变量，$y_k(n)$ 是对应的输出响应。

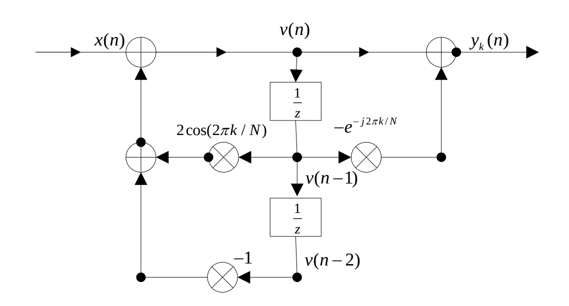


图 4 Goertzel 算法原理框图

Fig. 4 Goertzel algorithm block diagram

3.2 Goertzel 算法性能分析

对 DTMF 信号检测之前，需要将模拟信号转换成数字信号。从表 1 可知，DTMF 信号最高频率为 $f_{max}=1633Hz$，采用 $f_s=8000Hz$ 的采用频率显然满足奈奎斯特采样定理的要求，然后对采样后的数字信号利用 Goertzel 算法求出相应频率的幅值信息，即

$$|X(k)|^2=|y_k(N)|^2=\left|v(N)-e^{-j\frac{2\pi k}{N}}v(N-1)\right|^2 \tag{17}$$

而需要检测的 DTMF 信号频率 f_i 与离散傅里叶系数 k 存在着如下关系：

$$k=N\times\frac{f_i}{f_s} \tag{18}$$

其中：N 是采样点数，f_s 是采样频率。N 的选取取决于频率分辨率和采集 N 个采样点所需要的时间，N 取值越大，频率分辨率越高，采样所需时间也会相应增加。因此，需要在频率分辨率和采样时间这两个因素之间作出权衡。文献[16]中给出了 $N=205$ 是最佳值，这时的频率偏差保持在小于 1.5% 的范围内，符合国际电信联盟 ITU 对 DTMF 信号有效性的要求[17]，所以本文在实验时，取 $N=205$ 和 $f_s=8000$。利用上式可以求出 DTMF 信号的 8 个频率所对应的 k 值，如取 $f_i=697Hz$ 时，有

$$k=N\times\frac{f_i}{f_s}=205\times\frac{697}{8000}=17.8606 \tag{19}$$

由于 k 必须取整数，所以将其四舍五入取 $k=18$，DTMF 其他频率也可以通过上式相应计算得到，8 个 DTMF 频率对应 k 值如表 2 所示。

表 2　8 个 DTMF 频率对应的 k 值

Table 8　DTMF frequencies

基本频率(Hz)	计算 k 值	整数 k 值	绝对误差
697	17.861	18	0.139
770	19.731	20	0.269
852	21.833	22	0.167
941	24.113	24	0.113
1206	30.981	31	0.019
1306	34.235	34	0.235
1477	37.848	38	0.152
1633	41.864	42	0.154

由于 Goertzel 算法只关心 8 个频率的信息，相对于 FFT 算法计算量更小，在频域上的分辨率更好。如图 5 所示在 $N=205$ 和 $f_s=8000$ 条件下，数字“1”对应的 DTMF 信号识别情况，其中FFT算法将信号整个频域的信息都计算出来了，且需要找出两个幅度最大的频率值，而对应的 Goertzel 算法计算 8 个 DTMF 信号的幅度值，计算量减小，而且分辨率更高。如表 3 所示，对于 Goertzel 算法和 FFT 算法计算复杂度对比，在 $N=205$ 的条件下，Goertzel 的计算量相对于 FFT 算法减少了 60%。

4　算法实现

算法的流程如图 6 所示，具体的步骤为：

a)对输入的数字语音信号 $x(n)$ 进行预处理，其中包括滤波、分帧加窗使其去除大部分噪声信号。

b)信号分帧之后，每一帧记为 $s_i(n)$，$n=1, 2, \cdots, N$，N 为帧长，i 为帧数。以帧为单位计算语音信号的短时能量和短时过零率，分别为序列 $\{E_i\}$ 和 $\{Z_i\}$。

c)根据信号的短时能量和背景噪声的短时能量选取的阈值 T_1 和 T_2，进行短时能量的判别，确定语音段起止点的大致范围。

d)根据信号的平均过零率和噪声的平均过零率选取的阈值 T_3，进行短时过零率的判别，确定最终语音段的起止点。

e)循环步骤 c)和 d)，当满足双阈值判别时，将最近一次的能量序列作为语音信号端点的候选序列，直到遍历完所有的能量序列，得到最终的候选分段序列 $\{S_i\}$。

f)将分段序列 $\{S_i\}$ 对应的时域信号转换到频域，通过 Goertzel 算法进行 DTMF 信号的解析，找到每一段 DTMF 信号在频域内对应的频率对 (f_i, f_j)。

g)将所有的频率对在 DTMF 编码中查询，得到该信号最终的数字序列 $\{P_i\}$，算法结束。

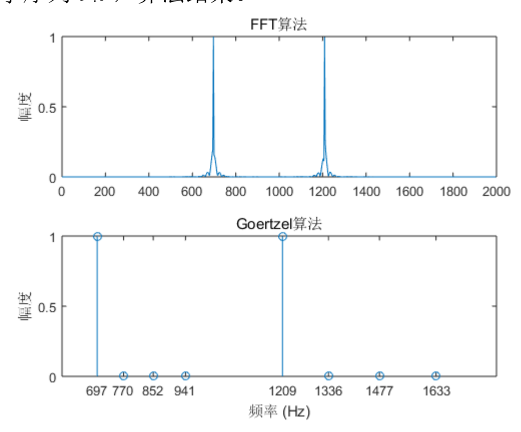


图 5　数字“1”的 FFT 变换与 Goertzel 变换

Fig. 5　FFT algorithm and Goertzel algorithm for number "1"

表 3　Goertzel 算法与 FFT 算法运算结果对比

Table 3　Comparison between Goertzel algorithm and FFT algorithm

	Goertzel 算法	FFT 算法
加法次数	$3N\log_2 N$	$(2N+1)\times 8$
乘法次数	$2N\log_2 N$	$(N+1)\times 8$
总运算次数	$5N\log_2 N$	$24N+24$
$N=205$	4944	7872

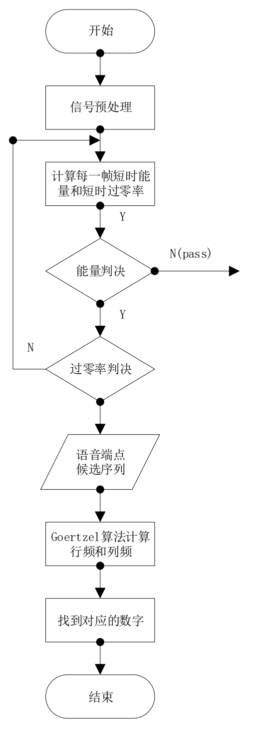


图 6　算法流程框图

Fig. 6　The algorithm flow diagram

5 实验结果与分析

5.1 实验环境

实验在杭州电子科技大学的第一教学楼实验室进行，首先在实验室环境下采集数据，分别按照 1 位、4 位、6 位和 11 位数字录取各 100 份录音数据，然后将录音数据在 Win10 系统下基于 MATLAB2016a 平台进行 DTMF 信号检测。其中录音采用两部 Android 手机相距 20 cm 进行录制，音频格式为 wav，信号采样率为 8 kHz，精度为 16 bit，窗函数采用汉明窗，帧长取 32 ms，帧移为 8 ms。

5.2 结果分析

首先将采集到的 300 份数据(4 位数字，6 位数字，11 位数字)，加入 6 种不同信噪比下的噪声，然后根据基于时域的端点检测算法进行分割。本文定义了端点分割的误差函数：

$$p_i = \frac{T_i}{G_i} \tag{20}$$

其中：p_i 为 i 位数字的分割误差，T_i 为 i 位数字经过基于时域的端点分割后的有效区间数，G_i 为 i 位数字有效区间的真实数(Ground Truth)。如图 7 所示，三种不同位数的数字在不同信噪比下的端点检测结果。可以看出 4 位数字和 6 位数字信号在信噪比很低的条件下，分割精确度在 80%以上，在高信噪比的条件下能达到 90%以上的分割准确率。而由于电话号码数字位数较长，在低信噪比的条件下，分割准确率相对 4 位数字和 6 位数字要低，但是也基本满足要求，分割准确率基本保持在 75%以上。

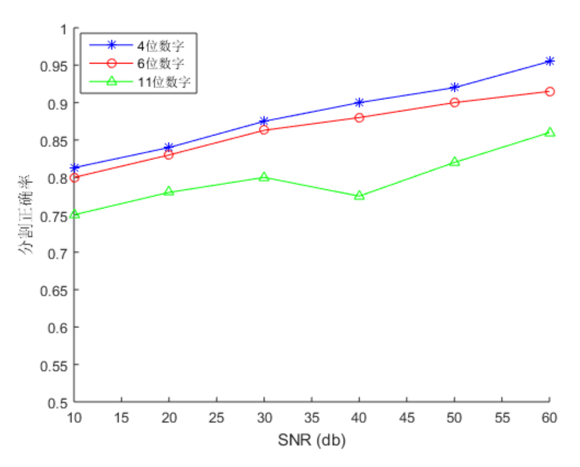


图 7 三种数字在不同信噪比下的分割误差

Fig. 7 Segmentation error of three numbers at different signal-to-noise ratios

最后，本文在 4 位、6 位和 11 位数字切割完成后，将所有有效区间的信号经过 Goertzel 算法进行分类，从数字“0”至数字“9”，检测 DTMF 信号所对应的数字是否正确。如图 8、9 所示，将采集的 1 位数字的录音数据以及分割后的 1 位数字的信号数据(取 500 份)，进行 Goertzel 分类算法的结果，Goertzel 算法的平均分类准确率为 80.36%，实验结果表明，该攻击方法是有效的。

6 结束语

本文总结了现通信环境下智能手机隐私安全方面的已有攻击种类，并提出了一种基于 DTMF 信号的智能手机外部攻击方法，通过对用户拨打电话、输入密码等操作的按键音录音，然后对按键的 DTMF 信号进行时域和频域的混合算法进行解析，获取用户按键信息。文中首先阐述了方案的具体实施过程；其次，介绍了基于短时能量和基于过零率的端点分割算法，本文将其二者进行结合，更好地利用了二者的特性，分割效果更好；然后，介绍了 Goertzel 算法的核心思想以及实现，将端点分割后的信号通过 Goertzel 算法进行分类；最后，对整个攻击方案的端点检测部分以及 DTMF 信号检测部分进行实验分析，证明了该方案的可行性和实用性。但是，由于文章篇幅和水平的限制，本文提出了攻击方案主要考虑在有噪声的环境下近距离对用户手机进行攻击，没有考虑到攻击的攻击角度、方向以及远距离等因素的影响，所以，下一步的研究方向将在这些方面继续深入研究。

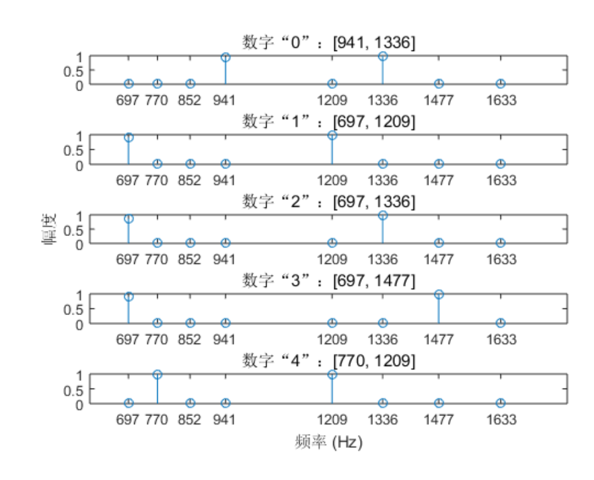


图 8 数字“0”到“4”的检测结果

Fig. 8 The detection results for numbers "0" to "4"

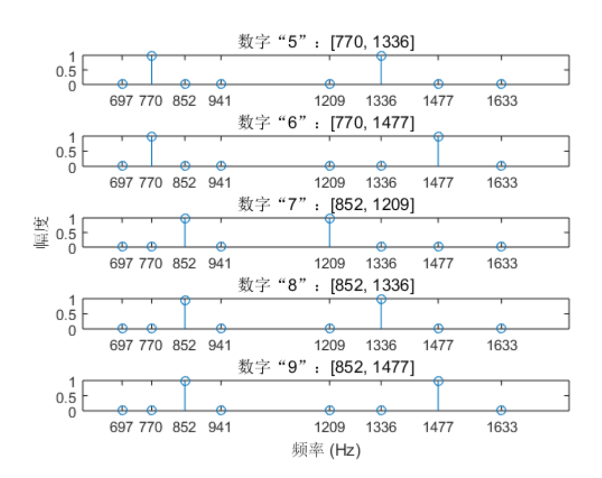


图 9 数字“5”到“9”的检测结果

Fig. 9 The detection results for numbers "5" to "9"

参考文献：

[1] Meeker M. Internet trends 2018 [EB/OL]. (2018-05-30) [2019-01-13]. https://www.kleinerperkins.com/perspectives/internet-trends-report-2018.

[2] Das A, Khan H U. Security behaviors of smartphone users [J]. Information & Computer Security, 2016, 24(1): 116.

[3] Schlegel R, Kapadia A, Wang X. Soundcomber: A Stealthy and Context-Aware Sound Trojanfor Smartphones [C]//Proc of Network and Distributed System Security Symposium. 2011.

[4] Drake J, Experts Found a Unicorn in the Heart of Android [EB/OL]. (2015-07-27) [2019-01-13]. http://blog.zimperium.com/experts-found-a-unicorn-in-the-heart-of-android/.

[5] Xing L, Bai X, Li T, *et al.* Unauthorized cross-app resource access on Mac OS X and iOS [EB/OL].(2015-09-07). https://sit.sice.indiana.edu/en/2015/09/07/xara-ccs/.

[6] Manikandan S P, Manimegalai R. Survey on mobile ad hoc network attacks and mitigation using routing protocols [J]. American Journal of Applied Sciences, 2012, 9(11): 1796-1801.

[7] Cheng N, Wang X O, Cheng W, *et al.* Characterizing privacy leakage of public WiFi networks for users on travel [C]//Proc of IEEE INFOCOM. Piscataway,NJ:IEEE Press, 2013.

[8] Akhuseyinoglu N B, Akhuseyinoglu K. AntiWare: an automated Android malware detection tool based on machine learning approach and official market metadata [C]//Proc of Ubiquitous Computing, Electronics & Mobile Communication Conference. Piscataway,NJ:IEEE Press, 2016: 1-7.

[9] Enck W, *et al.* TaintDroid: An Information-flow Tracking System for Realtime Privacy Monitoring on Smartphones [C]//Proc of Symposium on Operating Systems Design and Implementation. 2010.

[10] Khoula A H, Shah N, Shankarappa A N S. Smartphone's hotspot security issues and challenges [C]// Internet Technology & Secured Transactions. Piscataway,NJ:IEEE Press, 2017.

[11] 赵霞. 双音多频信号产生及解码的研究 [J]. 微电子学, 2001, 31(6): 418-421. (Zhao xia, The generation and detection of DTMF signals [J]. Microelectronics, 2001, 31(6): 418-421.)

[12] 张雅琪. 基于 MATLAB 的双音多频信号仿真 [J]. 信息与电脑:理论版, 2013(6): 129-130. (Zhang Yaqi. Dual-tone multi-frequency signal simulation based on MATLAB [J]. China Computer & Communication, 2013(6): 129-130.)

[13] 刘华平, 李昕, 徐柏龄,等. 语音信号端点检测方法综述及展望 [J]. 计算机应用研究, 2008, 25(8): 2278-2283. (Liu Huaping, Li Xin, Xu Boling, *et al.* Summary and survey of endpoint detection algorithm tor speech signals [J]. Application Research of Computers, 2008, 25(8): 2278-2283.)

[14] 韩立华, 王博, 段淑凤. 语音端点检测技术研究进展 [J]. 计算机应用研究, 2010, 27(4): 1220-1226. (Han Lihua, Wang Bo, Duan Shufeng. Development of voice activity detection technology [J]. Application Research of Computers, 2010, 27(4): 1220-1226.)

[15] Sysel P. Goertzel algorithm generalized to non-integer multiples of fundamental frequency [J]. Eurasip Journal on Advances in Signal Processing, 2012, 2012(1): 56.

[16] 金鑫春, 汪一鸣. Goertzel 算法下 DTMF 信号检测及参数优化 [J]. 现代电子技术, 2010, 33(6): 152-155. (Jin Xinchun, Wang Yiming, DTMF Signal Detection and Parameter Optimization Based on Goertzel Algorithm [J]. Modern Electronic Technique, 2010, 33(6): 152-155.)

[17] Technical features of push-button telephone sets [EB/OL]. (2003-02-10) [2019-01-13]. http://www.itu.int/rec/T-REC-Q. 23/en.